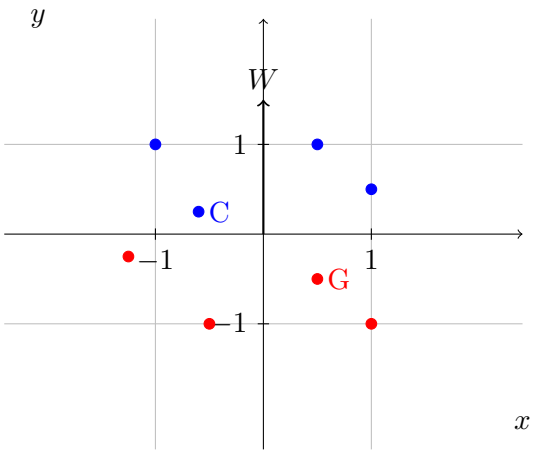
Deep Learning - Week 1

Common data for questions 1,2 and 3

In the figure shown below, the blue points belong to class 1 (positive class) and the red points belong to class 0 (negative class). Suppose that we use a perceptron model, with the weight vector w as shown in the figure, to separate these data points. We define the point belongs to class 1 if $\mathbf{w}^{\mathrm{T}}\mathbf{x} \geq 0$ else it belongs to class 0.



1. The points G and C will be classified as?

   Note: the notation $(G, 0)$ denotes the point G will be classified as class-0 and $(C, 1)$ denotes the point C will be classified as class-1

   (a) $(C, 0), (G, 0)$
   (b) $(C, 0), (G, 1)$
   (c) $(C, 1), (G, 1)$
   (d) $(C, 1), (G, 0)$

   Correct Answer: (d)
   Solution:

$$w = \begin{bmatrix} 0 \\ 1.5 \end{bmatrix},$$

$$x \in \begin{cases} 1 & \text{if } w^T x > 0 \\ 0 & \text{if } w^T x \leq 0 \end{cases}$$

For $C(-0.6, 0.2)$:

$$w^T x = \begin{bmatrix} 0 & 1.25 \end{bmatrix} \begin{bmatrix} -0.6 \\ 0.25 \end{bmatrix} = (0)(-0.6) + (1.25)(0.2) = 0.25$$

$\therefore (C, 1)$

For $G(0.5, -0.5)$:

$$w^T x = \begin{bmatrix} 0 & 1.25 \end{bmatrix} \begin{bmatrix} 0.5 \\ -0.5 \end{bmatrix} = (0)(0.5) + (1.25)(-0.5) = -0.625$$

$\therefore (G, 0)$

2. The statement that "there exists more than one decision lines that could separate these data points with zero error" is,

   (a) True

   (b) False

   Correct Answer: (a)
   Solution: The given statement is True.

   In the perceptron algorithm, when the data points are linearly separable, there can exist multiple hyperplanes (decision lines) that perfectly classify the data points with zero error. This is because a decision boundary depends on the orientation of the separating hyperplane and the margin around it, which can vary as long as it satisfies the linear separability condition.

   For example, in the graph provided, multiple lines can separate the red and blue data points such that all points are correctly classified. These decision boundaries can differ in slope and position while still achieving zero classification error. Hence, the solution is True.

3. Suppose that we multiply the weight vector w by -1. Then the same points G and C will be classified as?

   (a) $(C, 0), (G, 0)$

   (b) $(C, 0), (G, 1)$

   (c) $(C, 1), (G, 1)$

   (d) $(C, 1), (G, 0)$

   Correct Answer: (b)
   Solution: Simply multiply w by -1 and repeat the calculations from question 1.

4. Which of the following can be achieved using the perceptron algorithm in machine learning?

   (a) Grouping similar data points into clusters, such as organizing customers based on purchasing behavior.

   (b) Solving optimization problems, such as finding the maximum profit in a business scenario.

   (c) Classifying data, such as determining whether an email is spam or not.

   (d) Finding the shortest path in a graph, such as determining the quickest route between two cities.

   Correct Answer: (c)
   Solution: Perceptron can only classify, linearly separable data.

5. Consider the following table, where x_1 and x_2 are features and y is a label

x_1	x_2	y
0	0	1
0	1	1
1	0	1
1	1	0

Assume that the elements in $\mathbf{w}$ are initialized to zero and the perception learning algorithm is used to update the weights $\mathbf{w}$. If the learning algorithm runs for long enough iterations, then

(a) The algorithm never converges

(b) The algorithm converges (i.e., no further weight updates) after some iterations

(c) The classification error remains greater than zero

(d) The classification error becomes zero eventually

Correct Answer: (b),(d)
Solution: Since the data points are linearly separable, the algorithm converges, visualize it using a graphing tool.

6. We know from the lecture that the decision boundary learned by the perceptron is a line in $\mathbb{R}^2$. We also observed that it divides the entire space of $\mathbb{R}^2$ into two regions, suppose that the input vector $x \in \mathbb{R}^4$, then the perceptron decision boundary will divide the whole $\mathbb{R}^4$ space into how many regions?

(a) It depends on whether the data points are linearly separable or not.

(b) 3

(c) 4

(d) 2

(e) 5

Correct Answer: (d)
Solution: A line will become a hyperplane in $\mathbb{R}^4$ but still it will divide the region in 2 halves.

7. Choose the correct input-output pair for the given MP Neuron.

$$f(x) = \begin{cases} 1, & \text{if } x_1 + x_2 + x_3 < 2 \\ 0, & \text{otherwise} \end{cases}$$

(a) $y = 1$ for $(x_1, x_2, x_3) = (0, 0, 0)$

(b) $y = 0$ for $(x_1, x_2, x_3) = (0, 0, 1)$

(c) $y = 1$ for $(x_1, x_2, x_3) = (1, 0, 0)$

(d) $y = 1$ for $(x_1, x_2, x_3) = (1, 1, 1)$

(e) $y = 0$ for $(x_1, x_2, x_3) = (1, 0, 1)$

Correct Answer: (a),(c),(e)
Solution: Substituting values into the above expression and evaluating them yields the result.

8. Consider the following table, where x_1 and x_2 are features (packed into a single vector $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$) and y is a label:

x_1	x_2	y
0	0	0
0	1	1
1	0	1
1	1	1

Suppose that the perceptron model is used to classify the data points. Suppose further that the weights $\mathbf{w}$ are initialized to $\mathbf{w} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. The following rule is used for classification,

$$y = \begin{cases} 1 & \text{if } \mathbf{w}^{\mathrm{T}}\mathbf{x} > 0 \\ 0 & \text{if } \mathbf{w}^{\mathrm{T}}\mathbf{x} \leq 0 \end{cases}$$

The perceptron learning algorithm is used to update the weight vector $\mathbf{w}$. Then, how many times the weight vector $\mathbf{w}$ will get updated during the entire training process?

(a) 0

(b) 1

(c) 2

(d) Not possible to determine

Correct Answer: (a)
Solution: Upon computing $\mathbf{w}^{\mathrm{T}}\mathbf{x}$ for all data points with the initial weight values, all the points are correctly classified. Hence, update is not required.

9. Which of the following threshold values of MP neuron implements AND Boolean function? Assume that the number of inputs to the neuron is 3 and the neuron does not have any inhibitory inputs.

(a) 1

(b) 2

(c) 3

(d) 4

(e) 5

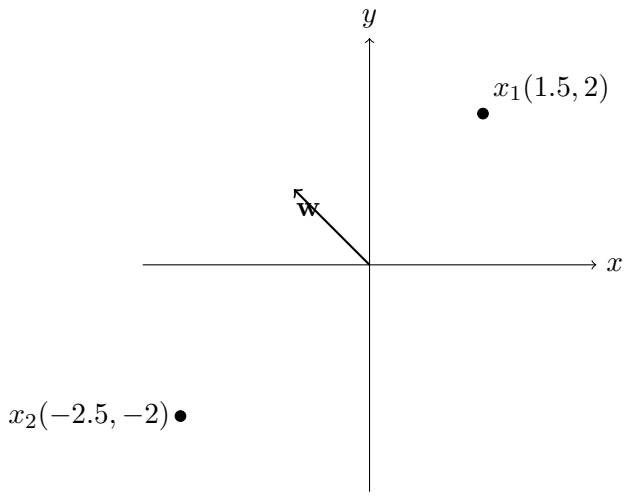
Correct Answer: (c)

Solution: suppose, we set $\theta = 4$, then summing all the input never exceeds 3, therefore, the neuron won't fire, And suppose we set $\theta < 3$ then it won't satisfy the AND operator.

10. Consider points shown in the picture. The vector $\mathbf{w} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$. As per this weight vector, the Perceptron algorithm will predict which classes for the data points x_1 and x_2.

NOTE:

$$y = \begin{cases} 1 & \text{if } \mathbf{w}^{\mathrm{T}}\mathbf{x} > 0 \\ -1 & \text{if } \mathbf{w}^{\mathrm{T}}\mathbf{x} \leq 0 \end{cases}$$



(a) $x_1 = -1$

(b) $x_1 = 1$

(c) $x_2 = -1$

(d) $x_2 = 1$

Correct Answer: (b),(d)

Solution: The decision boundary is $\mathbf{w}^{\mathrm{T}}\mathbf{x} = 0$. Hence for $\mathbf{w} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$, anything in the direction of $\mathbf{w}$ will have $\mathbf{w}^{\mathrm{T}}\mathbf{x} > 0$ and will get labelled 1.